UNIVERSITY OF ILLINOIS

May 8 19 85

THIS IS TO CERTIFY THAT THE THESIS PREPARED UNDER MY SUPERVISION BY

TERRY RONALD COLEY

ENTITLED DESIGN OF A MOLECULAR GRAPHICS PROGRAM WITH NOTES ON ITS APPLICATION TO CYTOCHROME C REDOX REACTIONS

IS APPROVED BY ME AS FULFILLING THIS PART OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE IN CHEMISTRY

Robert A. Scott
Instructor in Charge

APPROVED:

HEAD OF DEPARTMENT OF Chemistry

O-1364

DESIGN OF A MOLECULAR GRAPHICS PROGRAM WITH NOTES ON ITS APPLICATION TO CYTOCHROME C REDOX REACTIONS

by
TERRY RONALD COLEY

THESIS

for the
DEGREE OF BACHELOR OF SCIENCE
in
CHEMISTRY

College of Liberal Arts and Sciences
University of Illinois
Urbana, Illinois

1985

ACKNOWLEDGMENTS

My greatest thanks go to Dr. Robert Scott for the experience and freedom he has allowed me these past three years.

I also appreciate the friendship and assistance I have received from members of the Scott group.

Special thanks go to Laurie Carlson and my roommates for their understanding and assistance. Thanks also to Jean Kula for her assistance.

Congratulations to my fellow seniors in the Scott group who will graduate with me: Jean Kula and Dorothy Swain.

TABLE OF CONTENTS

TABLE OF FIGURES

Part I. Molecular Graphics Program

I.1. Introduction

Part I of this thesis describes a molecular graphics program, called DISPLAY, which has been under development since September of 1984. DISPLAY provides the following general capabilities to a scientist studying reactions involving combinations of structurally characterized proteins and small molecules:

1) Interactive, three dimensional visualization and orientation of proteins and small molecules.

2) Intermolecular and intramolecular distance calculations based on the selected orientation of objects.

3) Ability to function with inexpensive display hardware and hardcopy plotters.

Visualization is a useful tool when one is trying to understand the physical basis for a complex molecule's reactivity. For instance, when studying the interaction of a small molecule with a protein, one might postulate a theoretical expression for the rate of reaction which depends on spherical symmetry of the protein (1). This is obviously an approximation, but it is difficult to avoid without more detailed information about the steric interaction of the protein and small molecule. Furthermore, if the reactants are charged, an asymmetrical charge distribution on the large molecule may play an important

role in determining the rate of reaction by electrostatically favoring a positional orientation which brings the small molecule into contact with the protein reactive site. An x-ray structure of the protein is probably not sufficient by itself to easily provide information about sterically and electrostatically likely reactive sites. With three dimensional visual information, however, the investigator can much more easily suggest important protein reaction site(s) and attempt to relate these to a theoretical treatment of reaction rates. If the theory involves electrostatics, local potentials may well need to be computed often to test proposed reactive conformations with the theory. The inter- and intramolecular distance calculator in DISPLAY can provide the necessary radii from any reaction conformation selected. Without a program such as DISPLAY, calculations (especially the intermolecular ones) would be extremely tedious to obtain from raw crystallographic information, or even orthogonal angstrom coordinates. More will be mentioned about suggested applications of DISPLAY in section I.3 and Part II of this text.

DISPLAY was written on a VAX 11/780 running under VMS operating system and using the C language (2). C was chosen over FORTRAN primarily because of its superior control structure. This enables the programmer to handle complex concepts in a carefully structured and modular fashion. Furthermore, C and FORTRAN functions (subroutines) can be linked to the same executable program. DISPLAY uses FORTRAN routines from the System Plot Package (3) to produce metacode plot files

which are later translated for output to a Houston Instrument HI-PLOT DMP-7 plotter (4). Use of the metacode translator will be described in section I.5.3. Details of the source code layout, compilation, and linking of DISPLAY are described in section I.4.

DISPLAY was written to drive a VT-100 terminal equipped with a Selanar Graphics-100 board (5). This has a black and white vector graphics emulator with adequate resolution (1225, 240) to display fairly complex structures. The alpha carbon chain and heme of cytochrome c form a good stereo view on the screen. More complex view are available on the Houston plotter, and provisions are made for multicolor plotting.

I.2. Use of the DISPLAY Program

DISPLAY is primarily menu driven. The general flow followed by the user is to create various entities which may be displayed, choose which ones to display, then display and orient the objects on the screen. When a good view is obtained, the entities are ready for plotting or distance taking. During the course of running, DISPLAY uses and creates several input and output files. The following sections describe these files and how to use DISPLAY.

I.2.1. Input and Output Files

DISPLAY operates with seven types of files:

1) Brookhaven Protein Data Bank Files

   input, any file name

2) Small Molecule Coordinate Files

   input, *.VTS recommended file name

3) Protein Group Coordinate Files

   input/output, DIS*.VTA

4) Display Command Files

   input/output, DIS*.VTP

5) Distance Group Files

   input/output, DISDIST*.VTD

6) Distance Table Files

   output, DISDOUTO.VTD

7) Plot Output Files

output, IOP*

The meaning of each file type and the *'s will be described below.

I.2.1.1. Brookhaven Protein Data Bank Files

Brookhaven Protein Data Bank files contain fixed format records describing the structural information of various x-ray characterized proteins. Record formats are described in a Protein Data Bank Description (6). DISPLAY can read many of these files and produce protein entity files and display command files directly from this information.

I.2.1.2. Small Molecule Coordinate Files

Small molecule coordinate files are created by the user with an editor (such as EDT). The * may be any valid VMS file name and the .VTS extension is only a standardizing recommendation. These files contain lines with a name and the angstrom coordinates of each atom in the small molecule. If published x-ray crystallographic parameters for a molecule are available, the information required for the .VTS files may be obtained almost directly from the output of various programs available though the x-ray department. More on this is described in section I.5.1. The files must have 'carriage_return' carriage control as opposed to FORTRAN carriage control as output by the x-ray programs. Information on how to change file carriage control types is described in section

I.5.2. The editor (EDT) produces carriage_return output files unless the input file was FORTRAN carriage control, in which case the output of the editor is also FORTRAN carriage control. The format of small molecule coordinate lines is as follows:

atom_name x, y, z

Free format is used to make it easy to produce .VTS files with an editor. Only the first four characters of the atom_name are used, and the comma separators as shown are necessary. Any number of spaces may separate items, but the commas should come immediately after the x and y angstrom coordinates. Small molecule coordinate files contain no connectivity information and are used only as input for the distance calculator.

I.2.1.3. Protein Group Coordinate Files

Protein group coordinate files are similar to small molecule coordinate files in that they contain atom identifiers and angstrom coordinates with no connectivity information. The protein group coordinate files are written by the entity creation section of DISPLAY and contain exact copies of the ATOM or HETATM records read from a Brookhaven Protein Data Bank file. There are three classes of protein group coordinate files with the following names: DISBACK*.VTA, DISSIDE*.VTA and DISPROS*.VTA, where * is a digit 1-9. Each time a display entity is created, the coordinates of various groups go into the corresponding .VTA file as follows: atoms from selected portions of protein polypeptide backbone go into DISBACK*.VTA where * denotes the protein entity number being

created. Atoms of selected protein amino acid side chains, including the alpha C, go into the corresponding DISSIDE*.VTA file. Atoms of selected prosthetic groups, such as a heme, go into the corresponding DISPROS*.VTA file. Protein group coordinate files, like small molecule coordinate files, are read by the distance calculator.

1.2.1.4. Display Command Files

Display command files contain the list of display commands which produce a picture of an entity with the correct connectivity. The various display commands will be described in section I.2.3.2. There are two sets of display command files: DISPROT*.VTP and DISMOLE*.VTP, where * is a digit 1-9. The DISPROT*.VTP series contain protein display entities; the DISMOLE*.VTP series contain small molecule display entities. The entity number, *, of the DISPROT*.VTP files correspond to the entity numbers, *, of the protein group coordinate files. That is, the DISPROT*.VTP file contains the connectivity information of the corresponding DISBACK*.VTA, DISSIDE*.VTA, and DISPROS*.VTA files. One or two of the protein group coordinate files may be nonexistent, depending on how the entity was constructed.

I.2.1.5. Distance Group Files

Distance group files have the same format as protein group coordinate files; they are a collection of Brookhaven Protein Data Bank file ATOM or HETATM records. There are two

distance group files: DISDIST1.VTD and DISDIST2.VTD. These are created by the distance calculator by copying selected records from small molecule coordinate files and protein group coordinate files to either DISDIST1.VTD or DISDIST2.VTD. Small molecule coordinate records are formatted to match the protein group coordinates file records.

I.2.1.6. Distance Table Files

Distance table files are produced when the distance calculator reads the distance group files and writes tables of the distance between each record of coordinates in DISDIST2.VTD (containing distance group two) with respect to each atom in DISDIST1.VTD (distance group one). This will be discussed more in section I.2.3.3. The distance table file is always named DISDOUT0.VTD. If more than one distance calculation is performed, new versions of DISDOUT0.VTD are created.

I.2.1.7. Plot Output Files

Plot output files are created when the set plot enable display command (spe) is used. All output to the screen is then also routed to metacode files IOP*. Furthermore, titles for plotted output may be included in the plotter output. This will be discussed more in section I.2.3.2.

I.2.2. Creating Display Entities

This section describes how to create an entity to be displayed. Up to 18 entities may be directly accessed by

DISPLAY at one time: DISPROT*.VTP and DISMOLE*.VTP where the *'s are digits 1-9. Up to ten of these may be displayed together and manipulated entirely independently.

I.2.2.1. Creating a Protein Entity

Figure 1 shows the menu selections used to begin creation of a protein entity. A Brookhaven file name must be specified. If the file is nonexistent, an error message is displayed and the menu returned. Note that all figures of proteins in this text are of oxidized cytochrome c from tuna, Brookhaven file CYTCOX.ALB. An output file, DISPROT*.VTP must also be specified. Any digit 1-9 may be chosen for *, but if the .VTP file already exists, a new version number will be created (overwritten as far as DISPLAY is concerned). The user is also prompted for a protein name for the entity. This name will be placed in an object command (o) at the beginning of the active protein command display file and will be used to identify what the DISPROT*.VTP file contains. It will also be used to refer to the entity in the the display and orient portion of DISPLAY.

Essentially any portion of the protein may be selectively included in the .VTP file to be later displayed. The groups within the protein entity are divided into backbones, side chains, and prosthetic groups, as the last menu in figure 1 shows. Figures 2, 3 and 4 show the menus for creating groups of backbone, side chains, and prosthetic groups, respectively. For each of these types of group, a protein chain identifier

Figure 1: Menu selections used to begin creation of a protein entity.

```
1)  Set Up
2)  Display and Manipulate
D)  Done
> 1

        1)  Build Protein Entities (Brookhaven file)
        2)  Specify Active Display Files
        D)  Done
        > 1

Enter name of Brookhaven file > cytcox.ab

Protein Files available :

None.

Enter a display file number to use for output
(enter one digit only, 1 - 9) ? 1

Enter name of protein > cyt

                1)  Backbones
                2)  Side Chains
                3)  Prosthetic Groups
                D)  Done
                >
```

Figure 2: Menu selections used to create a group of protein backbone atoms.

```
                    1)  Backbones
                    2)  Side Chains
                    3)  Prosthetic Groups
                    D)  Done
                    > 1

Enter the chain identifier > O

                              1)  Alpha Carbon Chain
                              2)  Peptide Bond Chain
                              D)  Done with frame of Backbones
                              > 2

                                       1)  Whole Chain
                                       2)  Specific Sequence Numbers
                                       D)  Done
                                       > 2

Enter First residue number > 1
Enter Last  residue number > 17

Working . .


Do you wish to include carbonyl oxygens (Y/y = yes) ? n

                                       1)  Whole Chain
                                       2)  Specific Sequence Numbers
                                       D)  Done
                                       > d

                              1)  Alpha Carbon Chain
                              2)  Peptide Bond Chain
                              D)  Done with frame of Backbones
                              > d

                    1)  Backbones
                    2)  Side Chains
                    3)  Prosthetic Groups
                    D)  Done
                    >
```

Figure 3: Menu selections used to create a group of protein side chain atoms.

```
                    1)  Backbones
                    2)  Side Chains
                    3)  Prosthetic Groups
                    D)  Done
                    > 2

Enter the chain identifier > O

                              1)  All Types of Side Chain
                              2)  Specify by Type
                              D)  Done with frame of Side Chains
                              > 2

Enter the residue identifier > LYS

                                    1)  Entire Chain
                                    2)  Specific Sequence Numbers
                                    D)  Done
                                    > 2

Enter First residue number > 1
Enter Last  residue number > 17

Working .

                                    1)  Entire Chain
                                    2)  Specific Sequence Numbers
                                    D)  Done
                                    > d

                              1)  All Types of Side Chain
                              2)  Specify by Type
                              D)  Done with frame of Side Chains
                              > d

                    1)  Backbones
                    2)  Side Chains
                    3)  Prosthetic Groups
                    D)  Done
                    >
```

Figure 4: Menu selections used to create a group of protein prosthetic group atoms.

```
                    1)  Backbones
                    2)  Side Chains
                    3)  Prosthetic Groups
                    D)  Done
                    > 3

Enter the molecule identifier > 0

                              1)  Select Prosthetic Group
                              D)  Done with frame of Prosthetic Groups
                              > 1

Enter prosthetic group identifier > HEM

Working

                              1)  Select Prosthetic Group
                              D)  Done with frame of Prosthetic Groups
                              > d

                    1)  Backbones
                    2)  Side Chains
                    3)  Prosthetic Groups
                    D)  Done
                    > d

          1)  Build Protein Entities (Brookhaven file)
          2)  Specify Active Display Files
          D)  Done
          > d

1)  Set Up
2)  Display and Manipulate
D)  Done
>
```

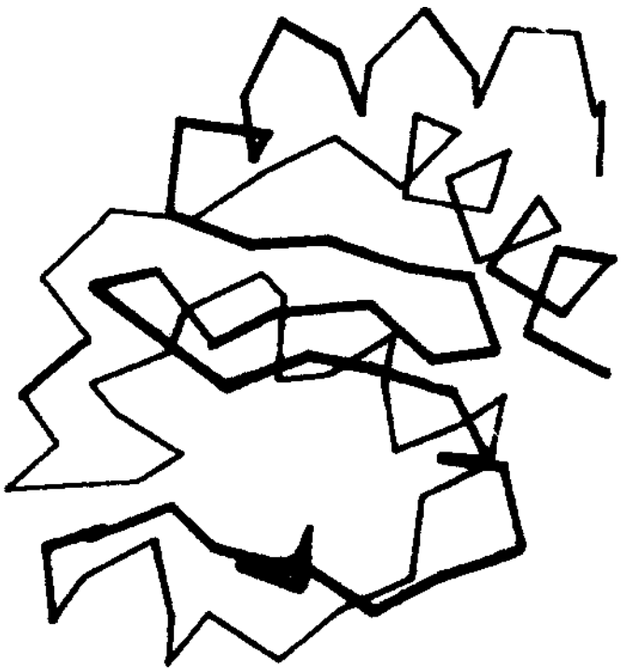
must be specified. For prosthetic groups, this is referred to as a molecule identifier. In each case the identifier is found in the Brookhaven file and is used to specify a chain in proteins with more than one chain, or a molecule in structures with more than one conformation.

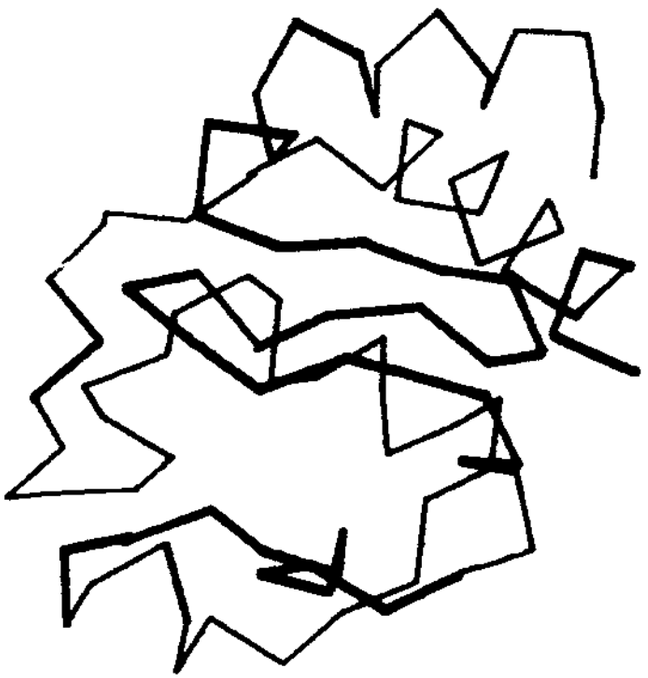
As indicated in figure 2, backbones can contain the full polypeptide chain or just the alpha carbons. Furthermore, any portion of these chains may be referenced by a starting and ending residue number (which may be equal). If alpha carbons are specified, a line is drawn between adjacent alpha carbons. If the peptide bond chain is specified, lines are drawn between adjacent N, alpha C, carbonyl C, N, ... atoms. The user is also asked if the carbonyl C to carbonyl O bonds should be included. The difference in complexity of these three ways of showing a backbone is illustrated for the full backbone of cytochrome c and a fragment of the backbone in figures 5, 6 and 7.

Figure 3 shows the selections available for side chains. As indicated, all side chains within a specified range may be selected, or only those of a specific type of amino acid. If a specific amino acid is specified, an identifier must be given. These are three letter identifiers readable from the Brookhaven file directly or from the Protein Data Bank Description. Side chain drawings include the alpha carbons, but they do not connect with adjacent alpha carbons. This must be done explicitly with the backbone creator. Samples of various side chain selections are shown in figure 8. A selecton of two or

Figure 5: Top: Stereo view of the full alpha carbon chain of cytochrome c.

Bottom: Stereo view of the alpha carbons 1-17 of cytochrome c.






17
1

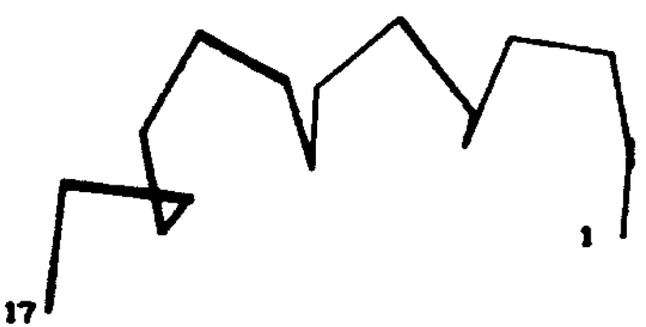
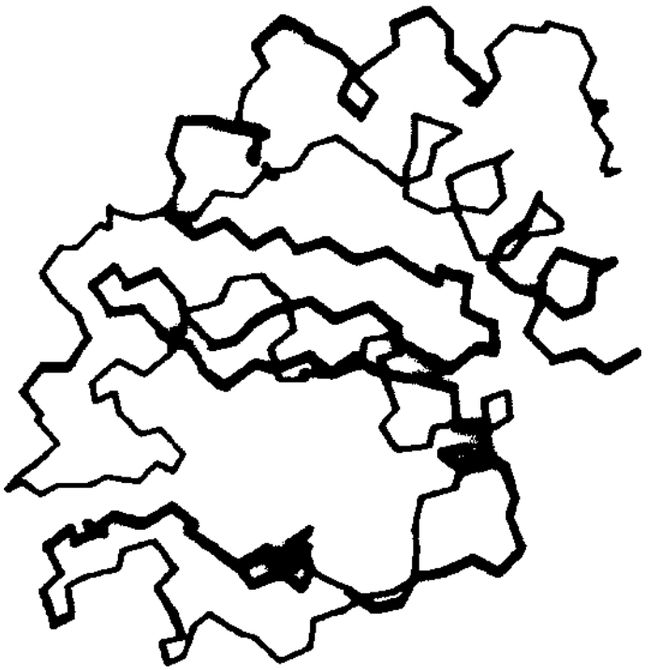
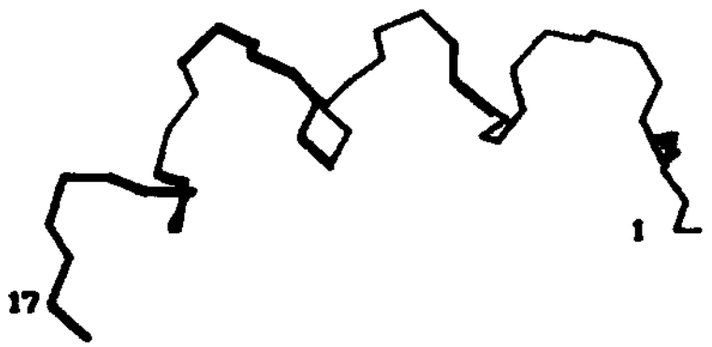

17
1

Figure 6: Top: Stereo view of all the polypeptide backbone linkages in cytochrome c.

Bottom: Stereo view of the polypeptide bond linkages for residues 1-17 in cytochrome c.




17
1

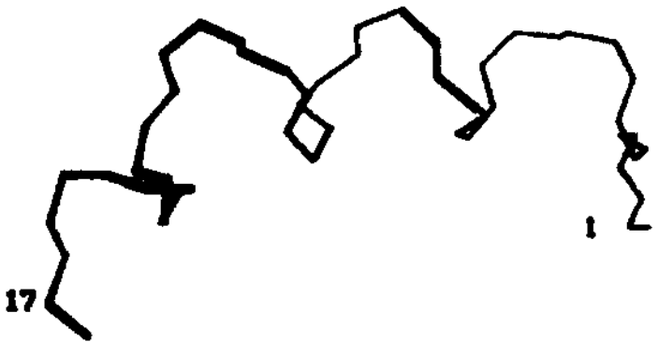
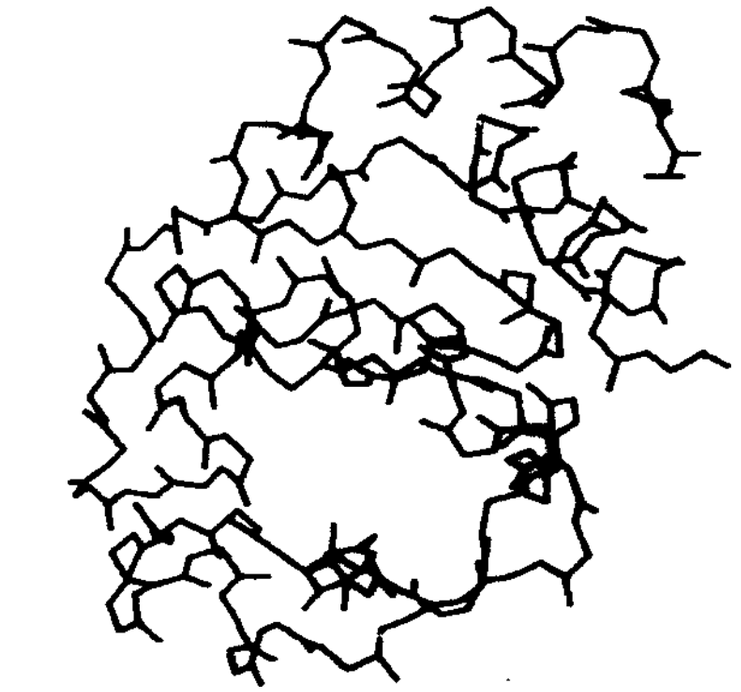
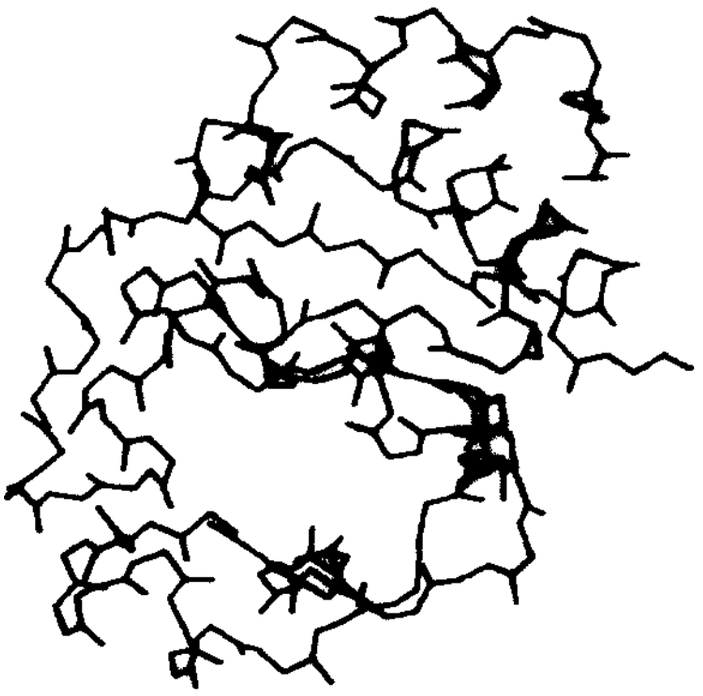
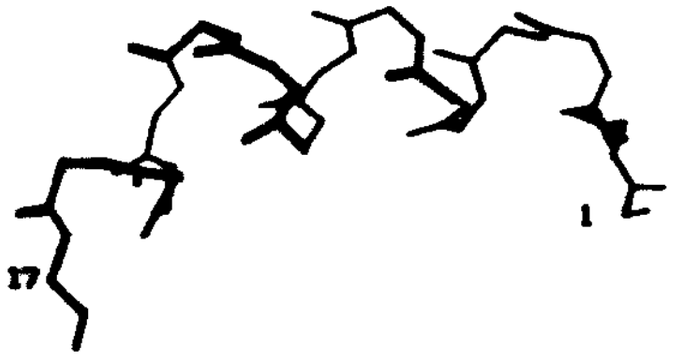

17
1

Figure 7: Top: Stereo view of all the polypeptide backbone linkages and carbonyl O's in cytochrome c.

Bottom: Stereo view of the polypeptide bond linkages and carbonyl O's for residues 1-17 in cytochrome c.






17
1

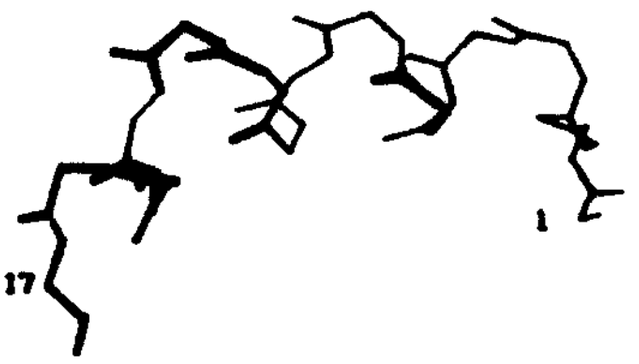
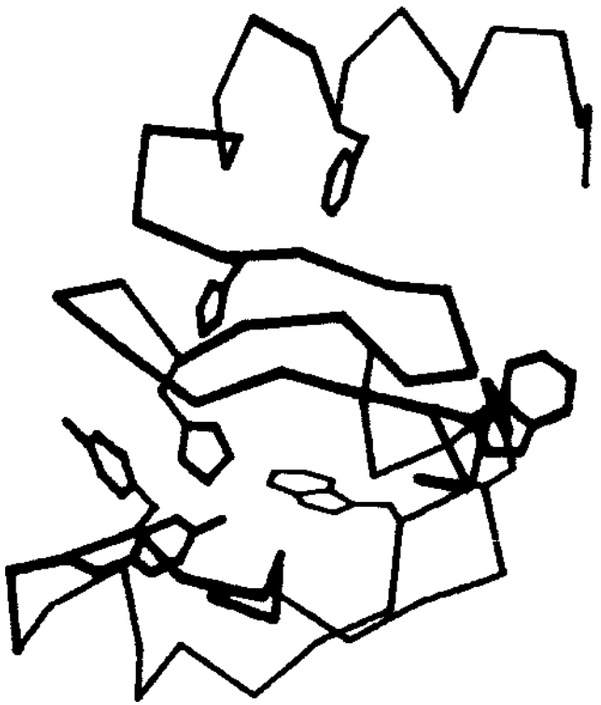
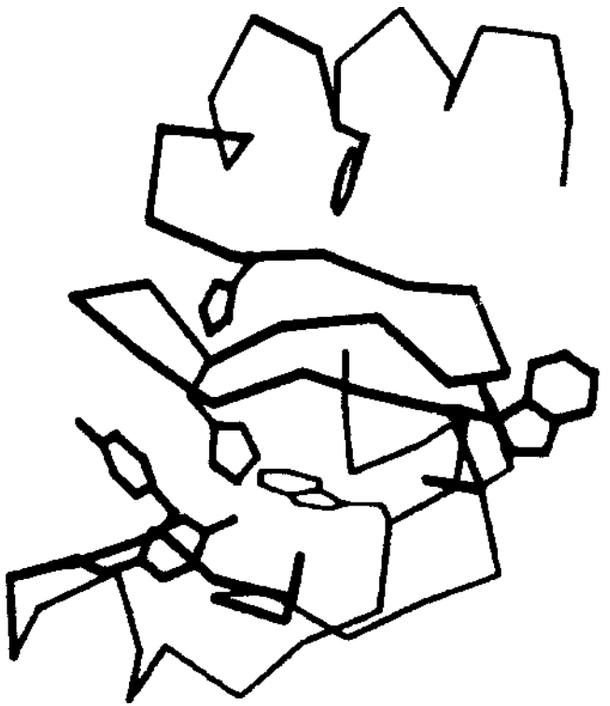

17
1

Figure 8: Top: Stereo view of all the side chains for residues 1-65 in cytochrome c.

Bottom: Stereo view of only the aromatic side chains for residues 1-65 in cytochrome c.





more specific types of side chain in a range of positions may be specified by repeated use of the appropriate menu selections.

Figure 4 shows the selection of a prosthetic group. This requires the use of a three letter identifier similar to the selection of specific side chains. Figure 9 shows the heme in cytochrome c.

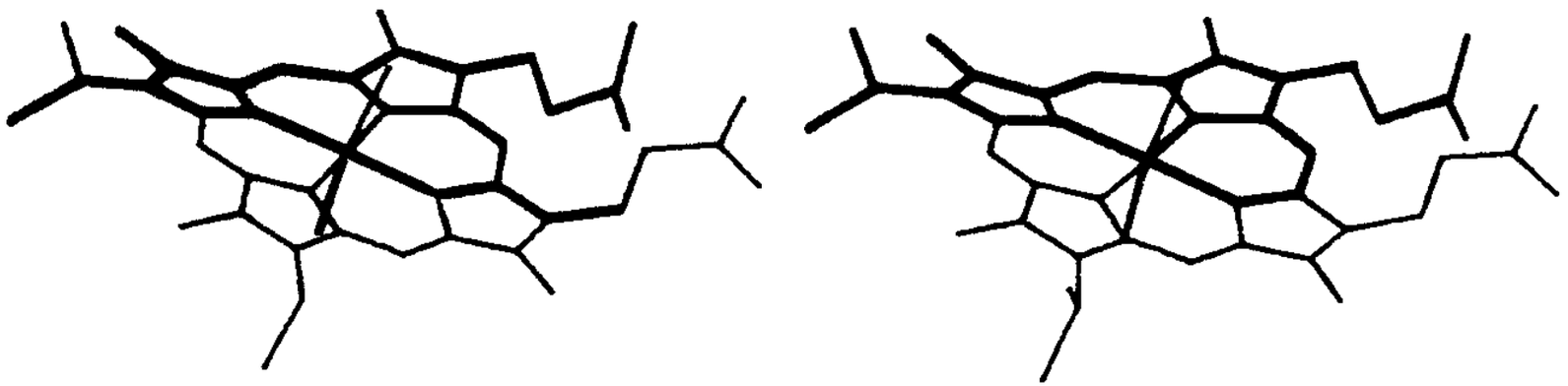
In each of the menus in figures 2, 3 and 4 the option "Done with a frame of ..." appears. If this is selected (as opposed to adding more to that group) a frame command (f) is added to the current display command file being created. Whenever a frame command occurs, the plotter pen stops temporarily and allows a change of pen colors. After the "Done with a frame of ..." option is selected, the same group creator (backbone, side chains, or prosthetic groups) may be chosen again and added to. This enable the user to have, for example, portions of the backbone in different colors or different types of side chain in different colors.

When all the groups have been created in a protein entity and the main menu is returned, a final frame command (F) and a stop command (S) are added to the protein display command file just created. This entity is now complete and ready for display.

I.2.2.2. Creating a Small Molecule Entity

Small molecules must be created outside DISPLAY by writing a coordinates file and a display command file with an editor such as EDT. The small molecule coordinates file is

Figure 9: Stereo view of the heme in cytochrome c.



described in section I.2.1.2. The display command file contains any list of valid display commands (discussed in I.2.3.2) with the following conditions. The first line must contain an object command (o) to specify an object name. The format is:

o object_name

where the o is in column 1 and at least one space separates the o from the object_name. The second command must be a window command (W) to specify the bounds of the molecule. The format is:

W xmin, ymin, xmax, ymax, zmax, zmin

where the W is in column 1 and at least one space separates the W from the first window specification. The first five numeric parameters must be separated by a comma immediately after the number followed by at least one space. The last command must be a stop command (S) with format:

S

where the S is column 1. The display command file should also contain one or more frame commands (F) if the plotter pen is to stop after drawing the small molecule or any part of it. One frame command is usually placed on the line before the S command.

The intervening display commands are usually a sequence of up pen (u), down pen (d), and move to absolute point commands (P) where the parameters of the P command are angstrom atomic coordinates. An initial translation or rotation may also be specified after the window command and before the pen

movement commands to place the small molecule in a desired initial orientation. Notes on obtaining atomic coordinates from x-ray crystallographic information are given in section I.5.1. The display commands are discussed in section I.2.3.2.

Once the small molecule display command file is created, it must be copied to one of the files DISMOLE*.VTP where * is a digit 1-9. The DISPLAY program can now access the .VTP file for display. Note that if the distance calculator is not used, a small molecule coordinates file does not need to be created.

I.2.3. Displaying and Manipulating Entities

This section describes how to specify active display command files, display and orient them, and how to use the distance calculator.

I.2.3.1. Specifying Active Files

The menu selections for specifying active files are shown in figure 10. A list of protein display command files appears first. The list shows what entity name is associated with the file and if it was previously active. At this point, the user enters a series of single digits each followed by a carriage return to specify the new active files. Only those specified will be active. A final carriage return with no file number will exit the activation procedure. The method of specifying active small molecule display command files is exactly analogous.

Figure 10: Specifying active files.

```
1)  Set Up
2)  Display and Manipulate
D)  Done
> 2

        1)  Display and Orient Views
        2)  Specify Active Display Files
        3)  Determine Distances
        D)  Done
        > 2

Protein Files available :

disprot1.vtp containing hp-65
disprot2.vtp containing hp-65a
disprot3.vtp containing heme

Enter file numbers for files you wish to be active display files
(enter last number only, (return) to stop) >
3


Small Molecule Files available :

None

        1)  Display and Orient Views
        2)  Specify Active Display Files
        3)  Determine Distances
        D)  Done
        >
```

1.2.3.2. Displaying and Orienting Entities

When the menu selection for displaying entities is chosen, each of the active display command files is interpreted and an image produced. The user is then prompted for a keyboard display command. The stop command (S) issued from the keyboard returns the user to the display and manipulate menu.

When the entities first appear, they are all in their own windows so the scales may not match. After interpreting all the display command files, however, DISPLAY sets the window of all display objects to that of the largest. If a clear graphics command (cg) is then issued, followed by an implement command (i), all objects will be redrawn to the same scale.

Figure 11 lists the online help pages which contain brief explanations of every display command. This list is obtained by typing a help command (h) or when any invalid keyboard command is issued. The formats shown indicate how parameters are specified in a display command file. Commands issued from the keyboard are not given in line format; only the keyword is entered, followed by a carriage return. DISPLAY then prompts the user for whatever command parameters are needed. Most of the keyboard parameter responses have defaults. The default values for the object transformation commands are 0.0; default parameters for most other commands are usually the current value of the parameter.

Most commands have the same effect when issued from the

Figure 11: Online display command help.

```
Object transformation commands ---

        All distances and angles are relative to the world coordinate system
        prior to global rotation.  All angles are in degrees.

c       = move center of rotation of object relative to its present center
C       = specify absolute center of rotation for active object
          format:  C x, y, z
c?      = show center of rotation of active object

r       = rotate object relative to its present rotation
R       = rotate object an absolute amount from original position
          format:  R theta_x, theta_y, theta_z
R?      = show rotation of active object

rg      = globally rotate world coordinate relative to its present rotation
Rg      = globally rotate world coordinate system an absolute amount
          format:  Rg theta_x, theta_y, theta_z
          The 'g' may be upper or lower case
Rg?     = show rotation of active object

More (y/n) ? y

t       = translate object relative to its present position
T       = translate object an absolute amount from original position
          format:  T x, y, z
th      = relative translation using the crosshairs
          If stereo viewing is enabled, th works on the right side view.
          (Valid from the keyboard only.)
T?      = show translation of active object

More (y/n) ? y

Object viewing parameter commands ---

        All points, vectors and angles are relative to the world coordinate
        system prior to global rotation.

E       = absolute set Eyepoint in world coordinates
          format:  E x, y, z
E?      = show Eyepoint coordinates of active object

N       = set view plane Normal vector
          format:  N x, y, z
N?      = show view plane Normal vector coordinates of active object

U       = set view Up vector, default (0, 1, 0)
          format:  U x, y, z
U?      = show view Up vector of active object

V       = set View reference point in world coordinates (default (0, 0, 0)
          format:  V x, y, z
V?      = show View reference point of active object

More (y/n) ? y

W       = set Window in world coordinates
          format:  W xmin, ymin, xmax, ymax, zfront (max), zback (min)
Wh      = set Window using the crosshairs
          If stereo viewing is enabled, Wh works on the right side view.
          (Valid from the keyboard only.)
W?      = show active window

vhg     = the viewing parameters for all objects are set to those of the
          active object.  This includes the Eyepoint, view plane Normal
          vector, view Up vector, and Window parameters.
vgi     = set viewing parameters to act individually again

More (y/n) ? y

Display characteristics commands ---

sd      = set physical label displacement (specified in pixels)
          format:  sd x_displacement_integer y_displacement_integer  (no comma)

sl      = set physical line thickening (more positive z coordinate lines are
          thickened by this number of pixels for depth cueing)
          format:  sl thickening_integer
```

```
spd      = set plotting disable
spe      = set plotting enable

ssd      = set stereo viewing disable
sse      = set stereo viewing enable (valid only from the keyboard)

More (y/n) ? y

st       = set physical tolerance (effective resolution)
           For example, physical tolerance of 1 gives full resolution because
           lines with endpoints at least one pixel apart are plotted.  Physical
           tolerance of 2 gives half resolution of end points, because only line
           endpoints separated by 2 or more pixels are plotted.

swd      = set words disable (ignore L commands)
swe      = set words enable (turn on L commands)

vp       = set viewport for screen.  If plotting is enabled, the plotter
           viewport parameters are required also.
           format:  vp xmin, xmax, ymin, ymax[, xmin_plot, xmax_plot, etc]
vp?      = show current viewports

More (y/n) ? y
Drawing commands ---

cg       = clear graphics screen
ct       = clear text

d        = pen down (valid from file only)

F        = Frame stops and asks for a title, halts plotter pen

I        = implement all display alterations since the last I command
i        = same as I command

l        = write label at specified coordinates (valid from file only)
L        = same as l command
           Requires 'set term/nowrap' to operate correctly.
           format   l x, y, z, label

o        = specify new active object
           format:  o string
o?       = list object names, show active object name

More (y/n) ? y

p        = relative move to point (valid from file only)
P        = absolute move to point (valid from file only)
           format:  P x, y, z, label

s        = start over from original file configuration (valid from keyboard only)

S        = when issued from the keyboard, returns to menu
           when issued from a file, stops file interpretation

u        = pen up (valid from file only)

Hit enter to continue...
```

keyboard as when executed from a display command file with the following exceptions:

1) The pen movement commands have no meaning when issued from the keyboard. These include up pen (u), down pen (d), pen move relative (p), and pen move to absolute point (P).

2) The stop command (S) issued from the keyboard returns to the DISPLAY menus. S issued from a command display file terminates interpretation of that file.

3) The start over command (s) should only be executed from the keyboard. This erases all the transformations (rotations, translations) executed since the display and orient section was first entered and reinterprets the active display command files.

4) The frame command (F) has no meaning when executed from the keyboard.

5) The crosshair commands (Wh and th) are keyboard commands only.

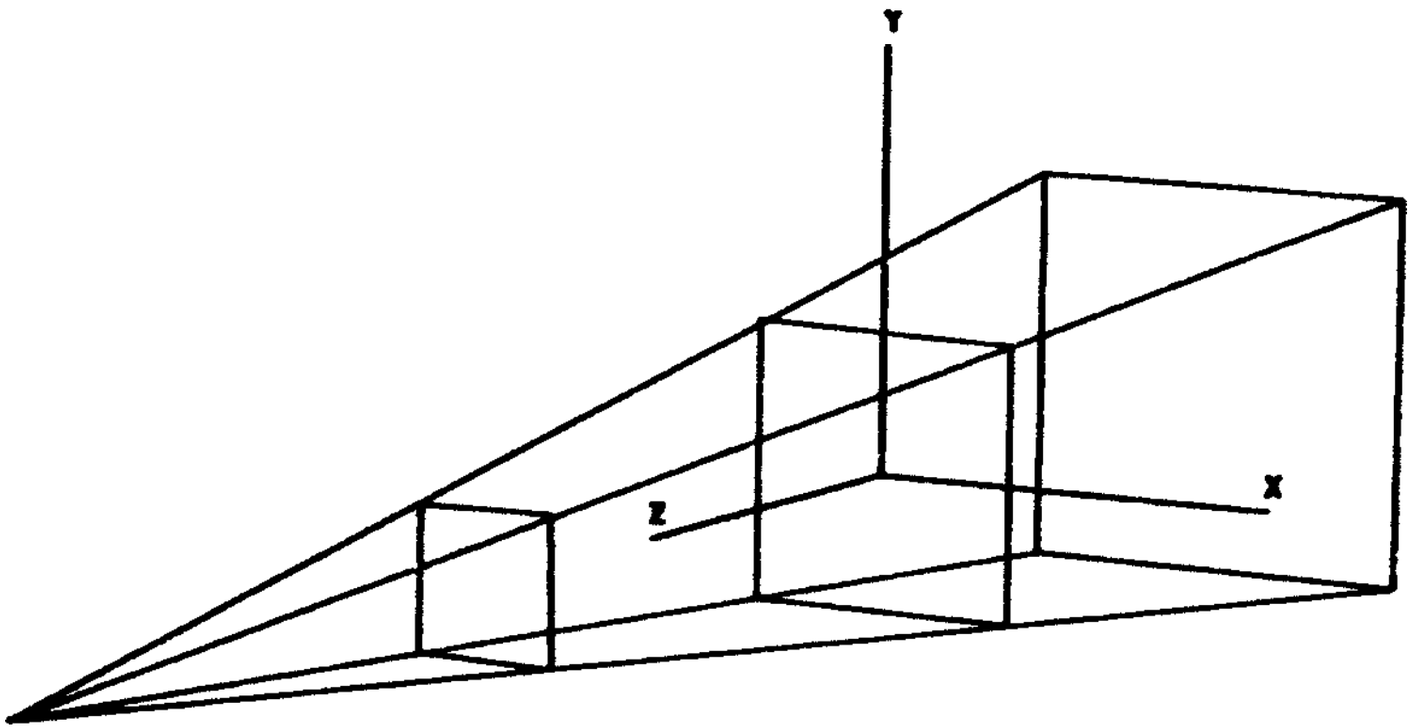
DISPLAY is capable of handling ten named display objects at once, completely independently. The object command (o) issued from the keyboard is used to specify an active object which the transformation commands and viewing parameter commands will apply to. The user is prompted for an object name. The o? command is used to display the names of all defined objects and show which one is active. DISPLAY assumes one named object per display command file as created in section I.2.2.

I.2.3.2.1. Object Viewing Parameter Commands

DISPLAY determines the largest window of all entities as specified in the active display command files, and then issues a view handle group command (vhg) internally before the user gets the viewing command prompt. This sets equal all the parameters affecting the vantage point from which an object is viewed. These parameters include the perspective eyepoint, the view reference point, the view plane normal, the view up vector and the window. The window command will be described in detail below. If any of the other viewing parameters need to be changed, refer to Foley and VanDam for a detailed description of their function (7). With the view handle group command in effect, any changes made to the viewing parameters of the active object are copied to all objects so that all are still viewed from the same vantage point relative to the world coordinate system. If objects need to be viewed on the same screen, but with different viewing parameters, use the view handle individual (vhi) command to set individual viewing. Use the o command to specify which object's parameters need to be changed, then alter the viewing parameters as needed. For the rest of this text, it is assumed that vhg is in effect.

Figure 12 shows the three dimensional viewing window. All points in world space (the space of the atom coordinates) are projected through the eyepoint which is the intersection of the lines joining the corners of the three planes shown. The points where these projectors intersect the center plane,

Figure 12: The three dimensional viewing window (view volume). The center plane is the view plane; the smaller and larger planes are the front and back z-clipping planes, respectively.


Y
X
Z

called the view plane, are mapped onto the physical screen coordinates or plotter surface. The first four parameters of the window command, xmin, ymin, xmax, ymax define the borders of the view plane. The smaller and larger planes are called the front and back clipping planes, respectively. The last two window parameters, zmax, zmin, specify the location of these planes, respectively. Only objects in world space which fall in the view volume defined by the front and back clipping planes and the projectors from the corners of the view plane will appear on the screen or plotter.

The window command may be used to zoom in on a particular section of a view, or cut out unwanted visual information. The Wh command allows the user to use the crosshairs of the GRAPHICS-100 screen to define new x and y window parameters. The front and back clipping planes remain the same. When using the Wh command, it is best to query the current window with a W? command and remember these values. Then, to return from a clipped view, the W command is used with the remembered values. Note that, as mentioned in figure 11, the Wh command can be used with the right side of a stereo view. Figure 13 shows a full stereo view and a close up window about the heme of cytochrome c.

I.2.3.2.2. Object Transformation Commands

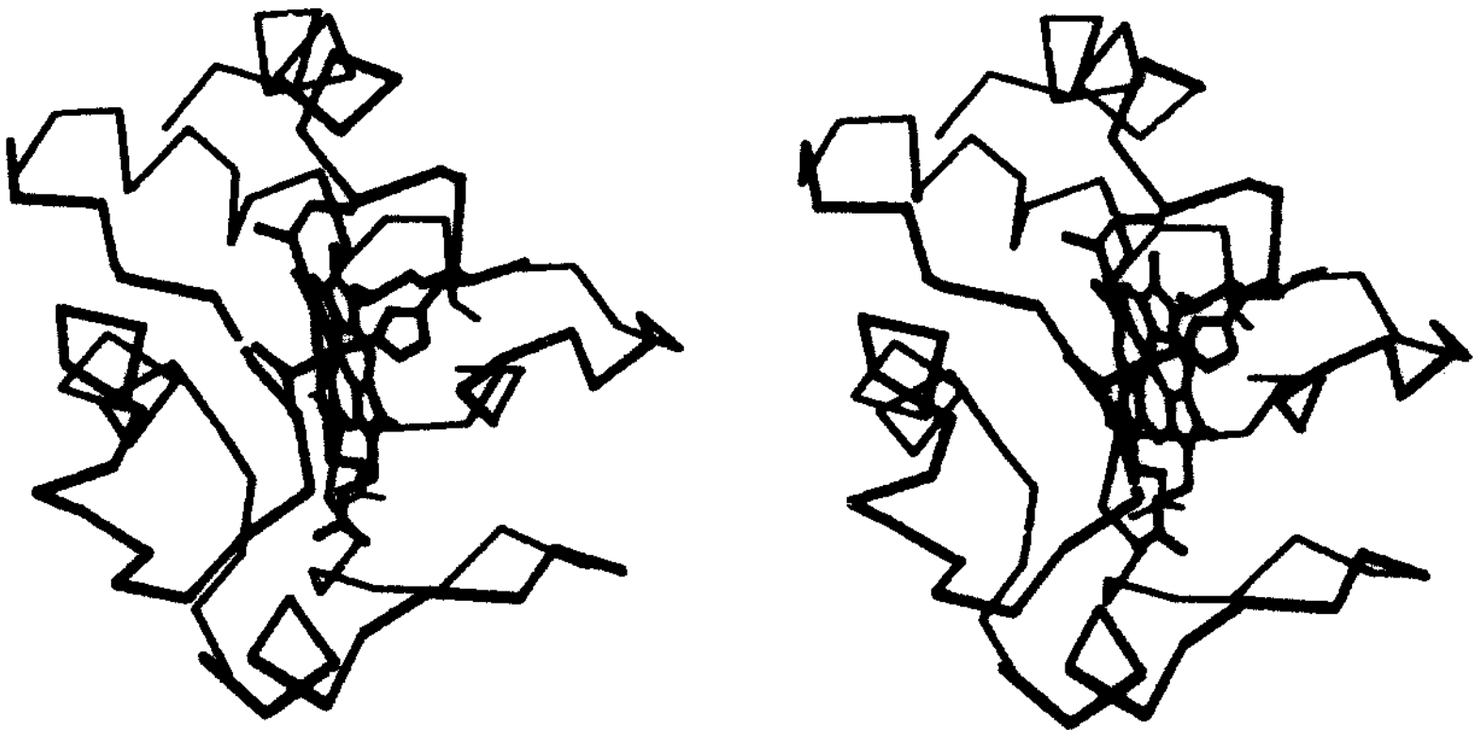
In addition to the viewing parameters, each object has associated with it orientation parameters which specify a translation from the object's originally defined world

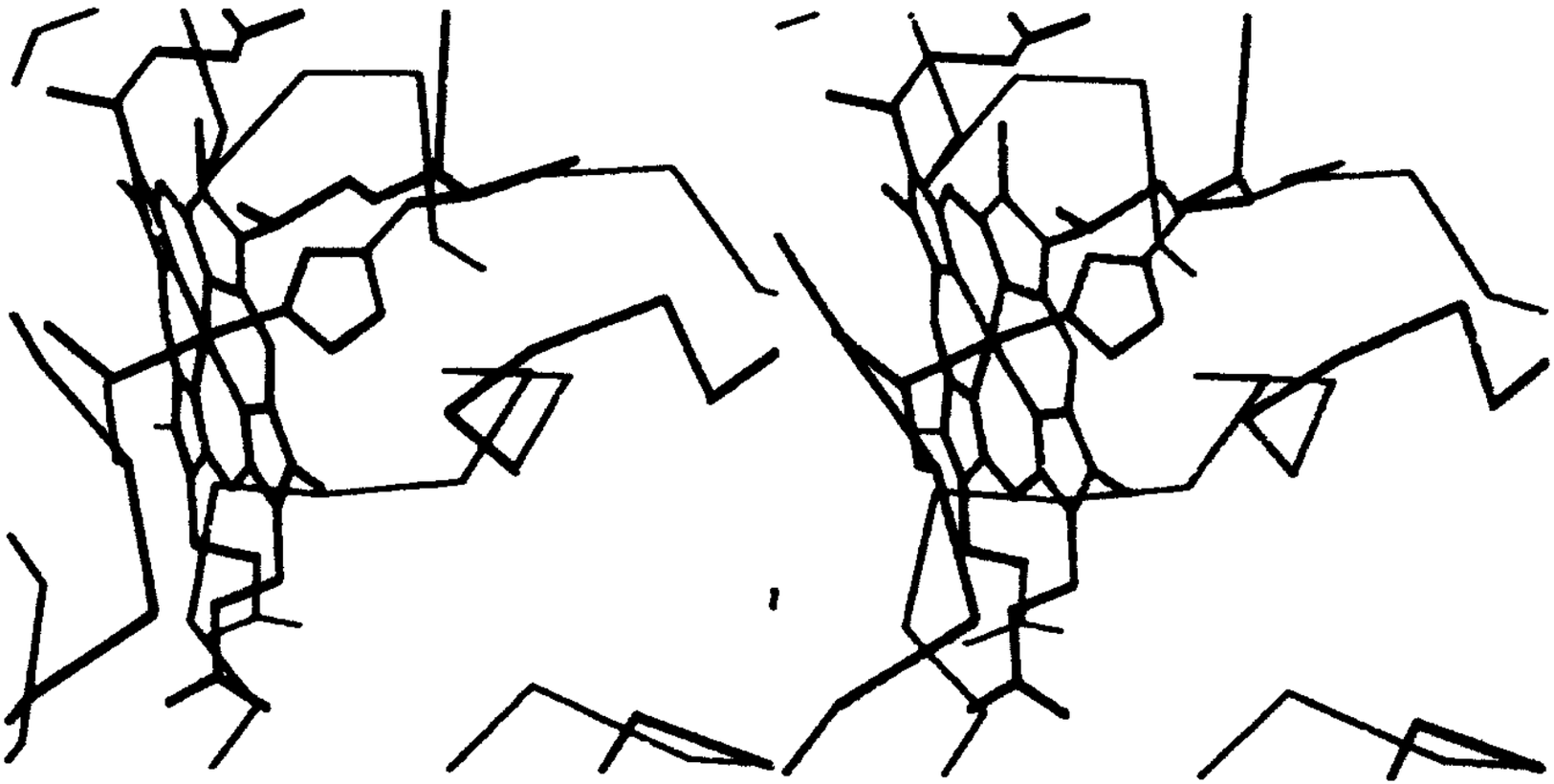
Figure 13: First Page:

Stereo view of cytochrome c alpha carbon backbone, cysteine 14, cysteine 17, histidine 18, methionine 80, plus the heme.

Second Page:

Stereo view of the same molecule in the same orientation with a zoom in on the heme and its attachments to the protein chain.





coordinates, followed by a rotation about a center of rotation defined in world coordinates. The translation is affected by the T and t commands. T performs an absolute translation of the object from its original coordinates by the specified amounts in world units. The t command performs a relative translation, adding the specified world units to the current translations in the x, y, and z-directions. An important difference between t and T is that the t command will relatively translate the object's center of rotation by the same amount as the translation.

The th command is also a relative translation, but it is performed using the crosshairs. When the th command is issued, the crosshairs appear and the user is prompted to place their intersection on the point where the center of rotation of the active object is to be translated to. This is an approximate method only. Furthermore, the th command changes only the x and y coordinates of the active object; the object remains in the same z-plane. The th command can be used with the right hand side of a stereo view.

The center of rotation may also be altered independently from the translation by using the C and c commands. These perform an absolute and relative translation of the center of rotation, respectively.

Rotation about the object's center of rotation is specified by the R and r commands. These perform absolute and relative rotations, respectively. The user is prompted for rotations about the three principal axes. These rotations are performed

optimal view of the synthesized molecule. The disadvantage of the current implementation of global rotation is that shifting the world coordinates orientation makes judging subsequent individual object translations and rotations more difficult because the new x, y and z directions must be kept in mind. Work is under way to provide global rotation and still allow subsequent object translations and rotations to be performed relative to the original world coordinate system.

I.2.3.2.3. Display Characteristics Commands

Display characteristics refer to how the given orientation and viewport will be presented. Most of the commands are self explanatory as seen in figure 11. A few notes are in order, however. The set tolerance command (st) is used to reduce plotting time for objects which may have too much detail in some area. Too much detail is defined as a change in world coordinates which is small enough not to produce a change in pen pixel position on the screen or plotter greater than or equal to some tolerance number of pixels. Normally, a view is selected such that this does not occur and the default value of the tolerance is one, as described in figure 11. If a view with a complex section is being plotted, time may be saved by setting the physical tolerance to a value greater than one. When this is done, pen movements smaller than the tolerance amount are not made. The distance of a pen move is measured from the last position which caused a move. This assures that a curve consisting of a series of small increments

will still be plotted, but with fewer effective increments.

The set physical line thickening command provides depth cuing. Lines are thickened by some fraction of the set line thickening parameter depending on their location between the front and back clipping planes. The thickening is determined linearly with z coordinates near the back clipping plane not thickened at all, and z coordinates near the front clipping plane thickened by the maximum number of pixels.

The set physical label displacement is used to offset the starting location of a point label (L command) by some number of character spaces so the label does not cover the point.

When plotting is enabled with the spe command, the user will be prompted for a plot title each time a frame (F) command is encountered in the display command file. The title may be blank (by typing carriage return) or a character string may be entered. This title will later be plotted near the top of the plotting surface. Also, the plotter pen will stop at these points during plotting and allow the user to change pens and decide if the next frame is to be plotted.

The stereo viewing option allows automatic production of stereo views. The user is asked if the views are to be viewed with a stereo viewer or with crossed eyes. The absolute angle separating the views is then specified, with a default value of six degrees. The default value is good for most views. A separation in centimeters is also requested, with a default of 6.5cm for a typical stereo viewer. Larger separations

may require a larger angular separation for maximum depth perception. The next two parameters are vertical offsets in centimeters. Using these offsets, stereo views of more than one orientation can easily be presented simultaneously. The only difference between stereo views for viewing with the eyes and those for viewing with viewers is the sign of the angular separation. This change is necessary to view the chirally correct molecule. Figure 14 demonstrates this. It is advantageous to learn how to view stereo views with the eyes because the separation and, therefore, the complexity can be increased. It is also useful to view an overhead projection of a stereo view.

The viewport command (vp) is used to specify what fractional portions of the screen and plotter surfaces are to be used. The full area is specified when the xmin, xmax and ymin, ymax parameters have the values 0.0, 1.0 and 0.0, 1.0, respectively.

I.2.3.2.4. Drawing Commands

The drawing commands have all been explained previously or in figure 11.

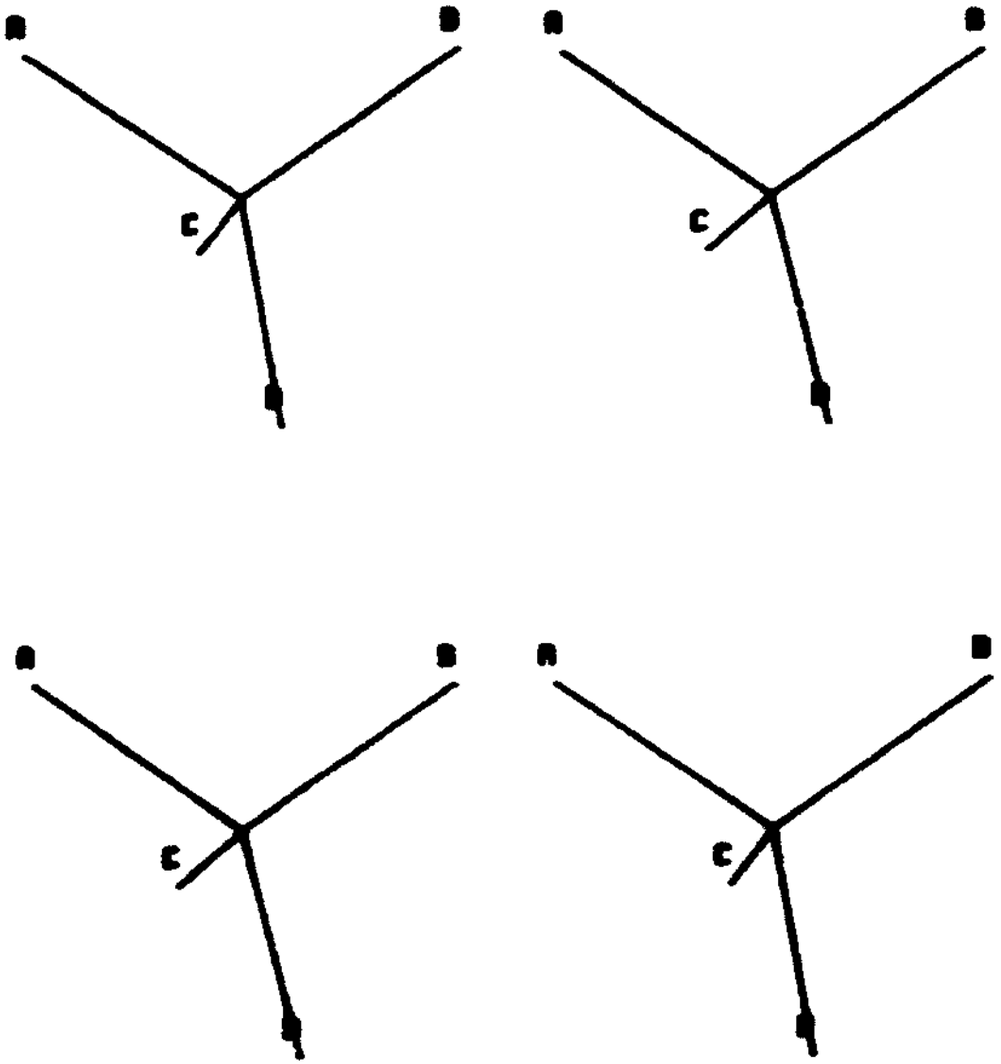
I.2.3.3. Distance Calculator

This section describes how to use the distance calculating facilities of DISPLAY. Before the distance calculator can be used, the display and orient option must be used. This sets up an object table and stores the object transformations. After the display and orient section is stopped (S), the distance

Figure 14: Top: Stereo view of a chiral molecule correct for crossed eyes.

Bottom: Stereo view of the same molecule correct for a stereo viewer.

Note that if the center to C bond is placed with C in back, A, B and D rotate clockwise in the chirality intended.


A
B
C
A
B
C
A
B
C
A
B
C

calculator may be selected. The distance calculator allows the user to specify two distance groups and then calculates every pair of distances between atoms. The distance groups may be intramolecular or intermolecular. A sample group selection is shown in figure 15. The distance groups selected are a small octahedral complex called 'oct' and the LYS 72 side chain of cytochrome c. Protein distance groups are specified exactly analogous to protein display entities (see section I.2.2.1). Small molecule distance groups are specified by giving a small molecule coordinates file name. Small molecule coordinates files are explained in section I.2.1.2. Figure 16 shows a stereo plot of the orientation chosen for these entities.

Distance tables are now ready to be output. The tables can be sent directly to the distance output file, DISDOUTO.VTD, or they may be seen on the screen also. Figure 17 shows the distance tables produced by the selected distance groups. Note that one table is produced for each atom in the second group. Therefore, the second distance group is usually specified as the smaller of the two groups, typically a small molecule. The minimum distance in each table is also indicated.

The tables include the residue name,chain or molecule identifier, and atom name for each atom of protein groups. The small molecule groups contain similar information with the first three letters of the object name replacing the residue name, a capital S replacing the molecule identifier, and the atom names coming from the small molecule coordinates

Figure 15: Distance group selection.

```
        1)  Display and Orient Views
        2)  Specify Active Display Files
        3)  Determine Distances
        D)  Done
        > 3

                1)  Select Distance Groups
                2)  Distance Groups Already Selected
                D)  Done
                > 1

Protein Files available :

disprot1.vtp containing hp-68
disprot2.vtp containing hp-65a
disprot3.vtp containing heme
disprot4.vtp containing cytheme
disprot5.vtp containing cytlys72

Small Molecule Files available :

dismole1.vtp containing window
dismole2.vtp containing chiral
dismole3.vtp containing Ru
dismole4.vtp containing oct

Specify Entity 1

Is entity 1 a protein or a small molecule (p/s) ? p

Enter the file number in which entity 1 resides > 5

                        1)  Backbone
                        2)  Side Chain
                        3)  Prosthetic Group
                        > 2

Enter the chain identifier > O

                        1)  All Types of Side Chain
                        2)  Specify by Type
                        D)  Done with group of Side Chains
                        > 1

                                1)  Entire Chain
                                2)  Specific Sequence Numbers
                                D)  Done
                                > 2

Enter First residue number > 72
Enter Last  residue number > 72

Working . . .

                                1)  Entire Chain
                                2)  Specific Sequence Numbers
                                D)  Done
```

```
                                    > d

                                    1)  All Types of Side Chain
                                    2)  Specify by Type
                                    D)  Done with group of Side Chains
                                    > d

Protein Files available :

disprot1.vtp containing hp-65
disprot2.vtp containing hp-65a
disprot3.vtp containing heme
disprot4.vtp containing cytheme
disprot5.vtp containing cytlys72

Small Molecule Files available :

dismole1.vtp containing window
dismole2.vtp containing chiral
dismole3.vtp containing Ru
dismole4.vtp containing eet

Specify Entity 2

Is entity 2 a protein or a small molecule (p/s) ? s

Enter the file number in which entity 2 resides > 4

Enter name of Small Molecule file > ru.vts

Working . . .

                                    1)  Output to Screen and File
                                    2)  Output to File only
```

Figure 16: Stereo view of an orientation of the alpha carbon backbone, LYS 72 portions of cytochrome c with an octahedral small molecule. This is the orientation used for the distance output of figure 17.

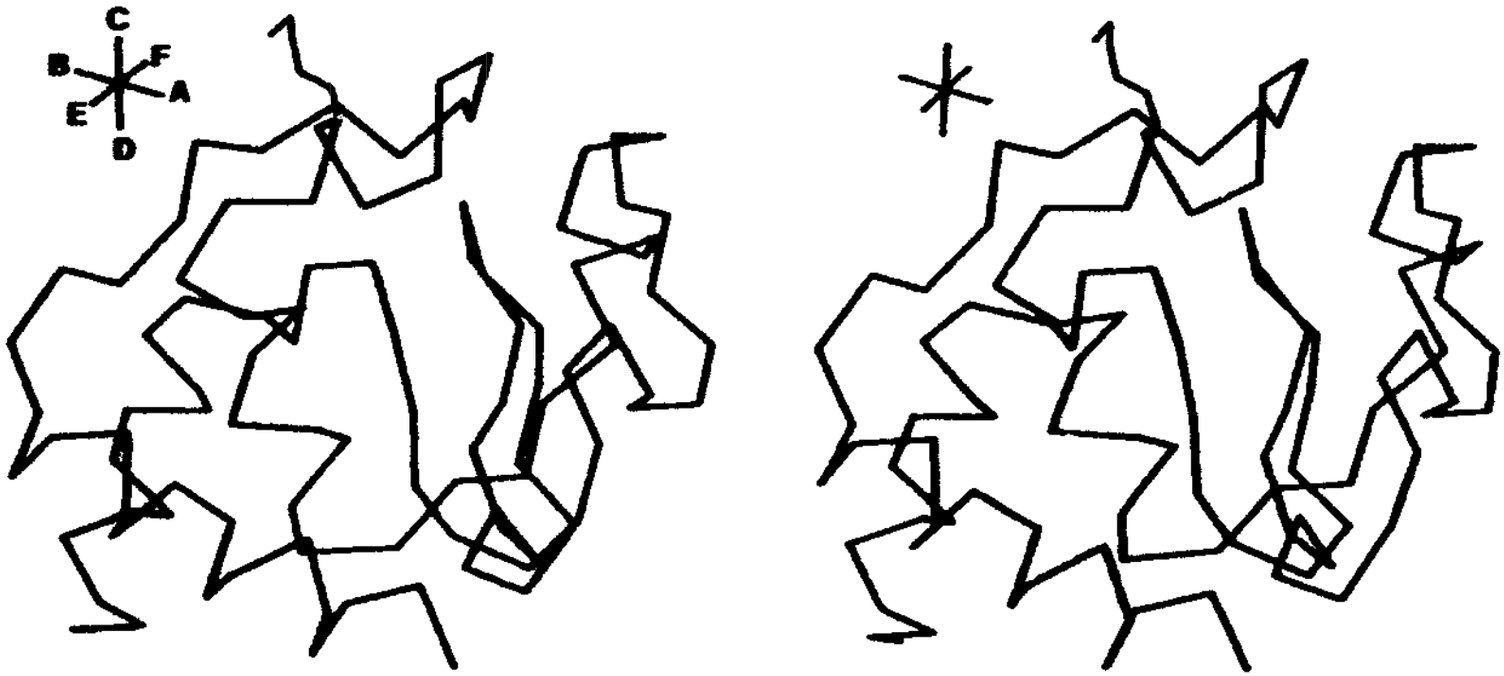

C
B
F
A
E
D

Figure 17: Distance tables for the distance groups selected in figure 15 with the orientation shown in figure 16.

```
oct    0  Ru     8     located at )  -9.676    10.342    10.000

--------------------- distance ----------------------------
LYS   72  CA     O        9.921       -1.004     9.205     5.317
LYS   72  CB     O        9.768       -1.454    10.523     4.729
LYS   72  CG     O        8.362       -2.737    11.042     5.386
LYS   72  CD     O        8.494       -3.023    12.553     5.205
LYS   72  CE     O        7.997       -3.069    13.268     4.574
LYS   72  NZ     O        6.735       -3.727    12.458     7.658


Closest distance :

LYS   72  NZ     O        6.735        1.917    -3.185    14.056
```

```
oct    0  A      S     located at )  -8.762    11.263     8.478

-------------------- distance ------------------------------
LYS   72  CA     O       8.627       -1.804     9.205     5.317
LYS   72  CB     O       8.247       -1.454    10.523     4.729
LYS   72  CG     O       6.774       -2.737    11.042     5.386
LYS   72  CD     O       6.732       -3.023    12.553     5.205
LYS   72  CE     O       6.330       -3.049    13.268     6.574
LYS   72  NZ     O       5.240       -3.727    12.458     7.658


Closest distance :

LYS   72  NZ     O       5.240        1.917    -3.185    14.056
```

```
oct    0  B      S     located at ) -10.589     9.421    11.522

--------------------- distance ------------------------------
LYS   72  CA     O       11.421       -1.004     9.305     5.317
LYS   72  CB     O       11.438       -1.484    10.533     4.729
LYS   72  CG     O       10.096       -2.737    11.042     5.384
LYS   72  CD     O       10.343       -3.035    12.553     5.305
LYS   72  CE     O        9.790       -3.049    13.248     6.574
LYS   72  NZ     O        8.461       -3.787    12.456     7.658


Closest distance

LYS   72  NZ     O        8.461        1.919    -3.185    14.056
```

```
oct    0  C      S     located at )  -9.586    12.037    11.073

--------------------- distance ------------------------------
LYS   72  CA     O       10.713       -1.004     9.308     5.317
LYS   72  CB     O       10.424       -1.454    10.533     4.729
LYS   72  CG     O        8.957       -2.737    11.042     5.386
LYS   72  CD     O        8.820       -3.023    12.553     5.285
LYS   72  CE     O        8.017       -3.069    13.248     6.574
LYS   72  NZ     O        6.794       -3.727    12.458     7.458


Closest distance :

LYS   72  NZ     O        6.794        1.917    -3.185    14.056
```

```
oct    0  D      S     located at )  -9.748     8.484     8.927

--------------------- distance ------------------------------
LYS   72  CA     O        9.498       -1.004     9.108     3.317
LYS   72  CB     O        9.496       -1.454    10.523     4.729
LYS   72  CG     O        8.235       -2.737    11.042     5.384
LYS   72  CD     O        8.431       -3.023    12.553     5.205
LYS   72  CE     O        8.464       -3.069    13.348     6.574
LYS   72  NZ     O        7.247       -3.727    12.458     7.488


Closest distance :

LYS   72  NZ     O        7.247        1.917    -3.185    14.056
```

```
oct    0  E      S     located at ) -11.453    10.900     9.272

--------------------- distance ------------------------------
LYS   72  CA     O       11.300       -1.004     9.205     5.317
LYS   72  CB     O       10.989       -1.454    10.533     4.729
LYS   72  CG     O        9.544       -2.737    11.042     5.384
LYS   72  CD     O        9.505       -3.033    12.553     5.205
LYS   72  CE     O        9.121       -3.069    13.248     6.574
LYS   72  NZ     O        8.045       -3.727    13.458     7.458


Closest distance :

LYS   72  NZ     O        8.045        1.917    -3.185    14.054
```

```
oct    0  F      S     located at )  -7.899     9.784    10.728

---------------------- distance ------------------------------
LYS   72  CA     O        8.786       -1.004     9.209     5.317
LYS   72  CB     O        8.836       -1.484    10.523     4.719
LYS   72  CG     O        7.534       -2.737    11.042     5.384
LYS   72  CD     O        7.871       -3.023    12.553     5.205
LYS   72  CE     O        7.242       -3.049    13.368     6.574
LYS   72  NZ     O        5.830       -3.727    12.458     7.458

Closest distance :

LYS   72  NZ     O        5.830        1.917    -3.185    14.054
```

file. The last three columns of numbers contain the transformed locations in world coordinates of each atom. The first column of number specifies the distance between that group 1 atom and the group 2 atom at the top of the table. The transformed coordinates can be useful in creating molecules from individual pieces of cyrstallographic structure. More on this is discussed in section I.3.1.

I.3. Example Uses of DISPLAY

The following are some possible ways to use DISPLAY. Sections I.3.1 to I.3.3 describe some techniques used to draw some pictures for a research proposal put forth by Sten Wallin of Dr. Robert Scott's group here at the University of Illinois. Part of Sten Wallin's research involves amide attachment of a pentaammine (isonicotinate) ruthenium (II) complex to terminal lysine nitrogens in cytochrome c. He then plans to study the rate of intramolecular electron transfer between the heme Fe and Ru centers.

I.3.1. Small Molecule Assembly

To produce a small molecule display entity for the Ru complex described above, the crystal structure for isonicotinic acid was obtained (8) and run through an x-ray department program to obtain angstrom coordinates. The x-ray programs are discussed in section I.5.1. Figure 18 shows the small molecule coordinates file and display command files which were created from the information provided by the x-ray department programs. Figure 19 shows the .VTS and .VTP files entered for a general octahedral complex with equal metal to ligand bond lengths of 2.0 angstroms. Displaying these results in figure 20. The x-ray programs produce coordinates for isonicotinc acid which place the molecular plane in the xy-plane. This can be verified

Figure 18: Top: Small molecule coordinates file for isonicotinic acid.

Bottom: Display command file for isonicotinic acid.

```
N1        2.3652,     0.3042,    -0.0041
C2        1.5073,     1.3242,    -0.0071
C3        0.1378,     1.1407,     0.0025
C4       -0.3652,    -0.1544,     0.0025
C5        0.5239,    -1.2116,     0.0053
C6        1.8837,    -0.9447,     0.0041
C7       -1.8496,    -0.4087,     0.0057
O1       -2.5626,     0.6707,    -0.0310
O2       -2.2939,    -1.5384,     0.0519
```

```
o  in
W  -5.0, -5.0, 5.0, 5.0, 5.0, -5.0
u
P   2.3651,  0.3042, -0.0041
d
P   1.5073,  1.3242, -0.0071
P   0.1378,  1.1407,  0.0025
P  -0.3652, -0.1544,  0.0025
P   0.5239, -1.2116,  0.0053
P   1.8837, -0.9447,  0.0041
P   2.3652,  0.3042, -0.0041
u
P  -0.3652, -0.1544,  0.0025
d
P  -1.8496, -0.4087,  0.0057
P  -2.2939, -1.5384,  0.0519
u
P  -1.8496, -0.4087,  0.0057
d
P  -2.5626,  0.6707, -0.0310
S
```

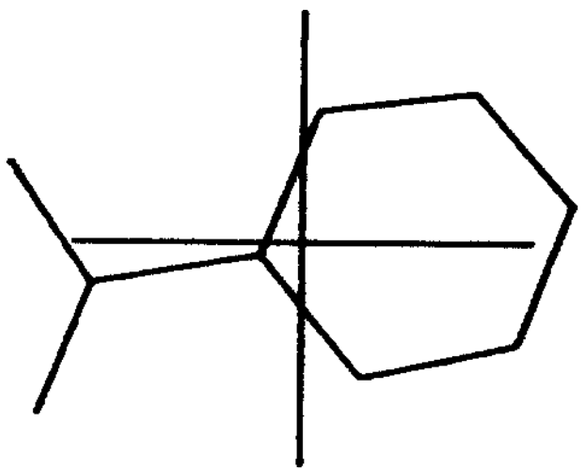
Figure 19: Top: Small molecule coordinates file for a symmetrical octahedral complex.

Bottom: Display command file for the above complex.

```
 Ru   0.00,   0.00,   0.00
 A    2.00,   0.00,   0.00
 B   -2.00,   0.00,   0.00
 C    0.00,   2.00,   0.00
 D    0.00,  -2.00,   0.00
 E    0.00,   0.00,  -2.00
 F    0.00,   0.00,   2.00
```

```
e Ru
W  -5.0, -5.0, 5.0, 5.0, 5.0, -5.0
u
P    2.00,   0.00,   0.00
d
P   -2.00,   0.00,   0.00
u
P    0.00,   2.00,   0.00
d
P    0.00,  -2.00,   0.00
u
P    0.00,   0.00,  -2.00
d
P    0.00,   0.00,   2.00
F
S
```

Figure 20: Isonicotinic acid and Ru complex displayed in their original positions.



by rotating the object named 'in' to find that a 90 degree rotation about the x or y-axis will make the molecule appear as a straight line, since it is then viewed edge on. The center of rotation of the Ru complex is the Ru atom which is located at the origin. Making sure 'Ru' is the active object (using the o command), an absolute translation of 2.0 angstroms in the x direction will leave the center of rotation at the origin and place one of the Ru ligands at the origin. This is shown is figure 21. Following this with a relative translation to the coordinates of the isonicotinic acid N1 (which is to be coordinated to the Ru) we obtain figure 22. The center of rotaton for the 'Ru' is now at the intermolecular joint. With a series of relative rotations about the z-axis, the C7-C4-N1-Ru-B atoms are made almost collinear, shown in figure 23. The Ru molecule is now rotated about the x-axis by 45 degrees to place the plane of the isonicotinic acid between the other Ru ligands. Stereo views of the resulting orientations are shown in figure 24. Since the isonicotinic acid has not moved, we can use the distance calculator to locate the Ru complex atoms relative to the original acid coordinates. Exiting the display section (S) and choosing the acid as distance group 2 and the Ru complex as group 1, we obtain the distance table in figure 25. This is the first table in the output file DISDOUT0.VTD and it contains the coordinates of each atom in the Ru complex. This information is used to edit a new small molecule coordinates file and a new display command file as shown in figure 26.

Figure 21: Acid and Ru complex displayed with the Ru complex translated 2.0 angstroms in the x direction.

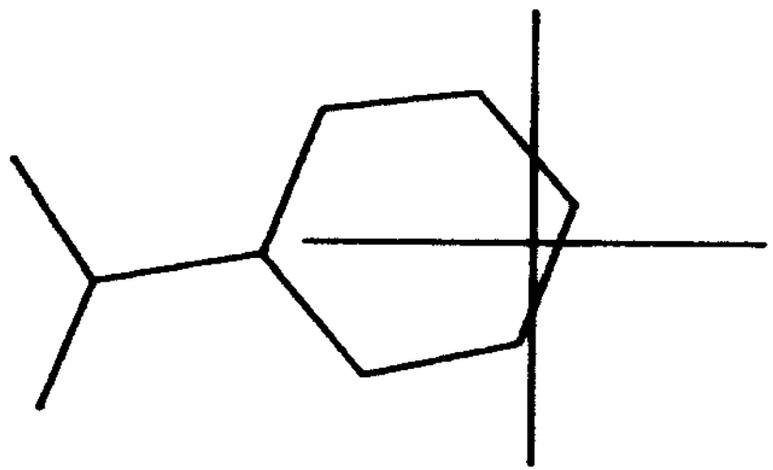


Figure 22: Acid and Ru complex displayed with the Ru complex subsequently translated to the N1 coordinates of the acid.

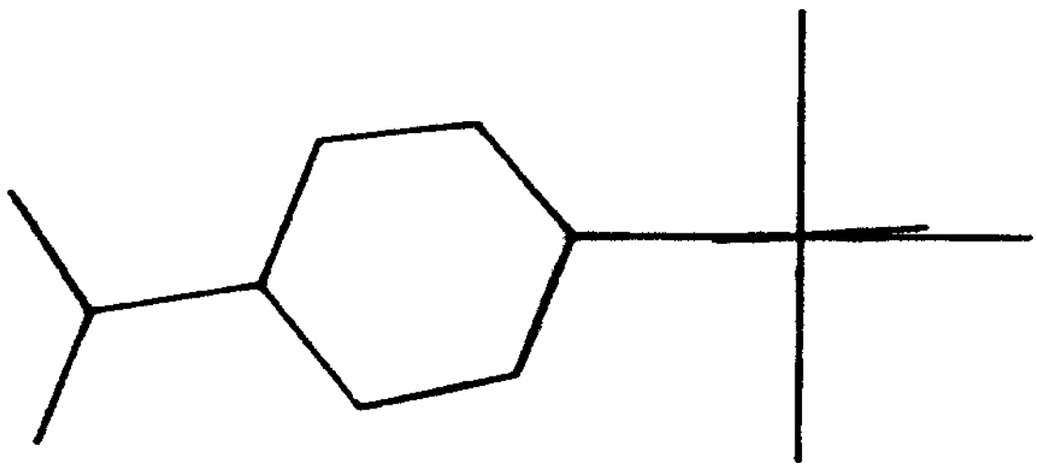


Figure 23: Acid and Ru complex displayed with the Ru complex rotated about the joint at acid N1 to make C7-C4-N1-Ru-B collinear.

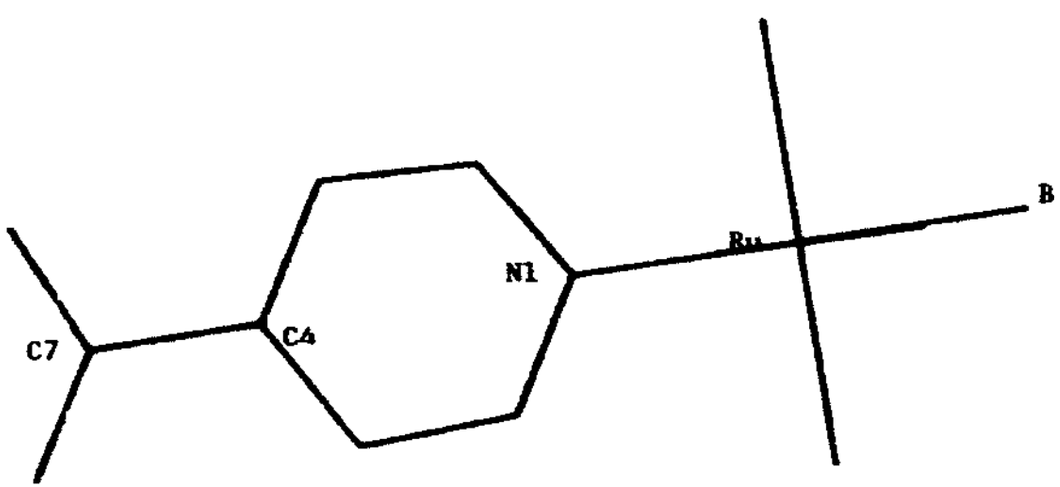

B
Ru
N1
C4
C7

Figure 24: Stereo view of the acid and Ru complex with the Ru complex subsequently rotated 45 degrees about the x-axis.

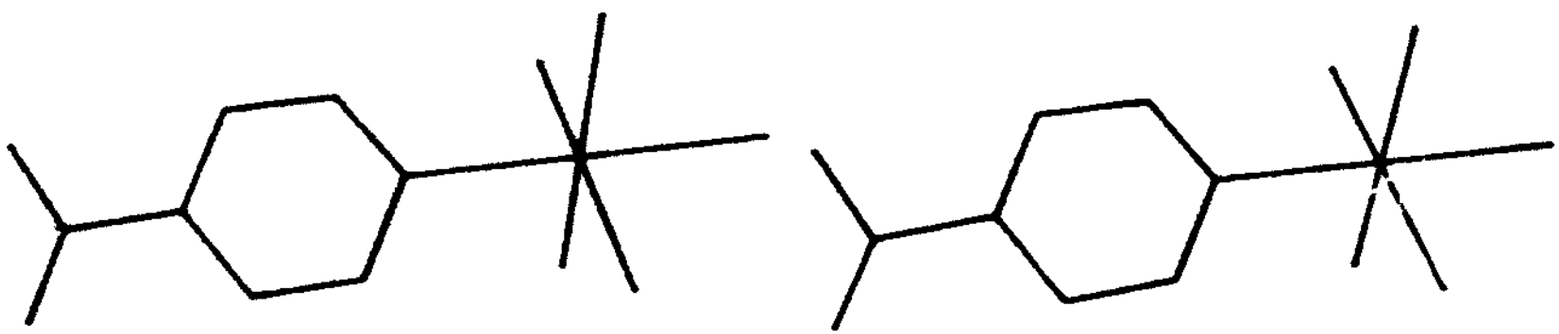


Figure 25: First distance table showing the coordinates of the Ru complex relative to the untranslated acid.

in	0	Ni	S	located at >	2.365	0.304	-0.004
				distance			
oct	0	Ru	S	2.000	4.338	0.535	0.227
oct	0	A	S	4.000	4.311	0.764	0.458
oct	0	B	S	0.000	2.365	0.304	-0.004
oct	0	C	S	2.829	4.012	1.930	1.622
oct	0	D	S	2.829	4.665	-0.860	-1.168
oct	0	E	S	2.829	4.338	1.949	-1.187
oct	0	F	S	2.828	4.338	-0.879	1.441

Closest distance :

oct	0	B	S	0.000	-2.000	0.000	0.000

Figure 26: First page:

A new small molecule coordinates file for the Ru - isonicotinate.

Second page:

A new display command file for the Ru - isonicotinate.

```
o  Ruin
W  -8.0, -8.0, 8.0, 8.0, 8.0, -8.0
u
P   2.3652,  0.3042, -0.0041
d
P   1.5073,  1.3262, -0.0071
P   0.1378,  1.1407,  0.0028
P  -0.3651, -0.1544,  0.0025
P   0.5239, -1.2116,  0.0053
P   1.8837, -0.9447,  0.0041
P   2.3652,  0.3042, -0.0041
u
P  -0.3651, -0.1544,  0.0025
d
P  -1.8496, -0.4087,  0.0057
P  -2.2939, -1.5384,  0.0519
u
P  -1.8496, -0.4087,  0.0057
d
P  -2.5624,  0.6707, -0.0310
u
P   4.311, 0.766, 0.458
d
P  2.365, 0.304, -0.004
u
P   4.012, 1.930, 1.622
d
P   4.665, -0.860, -1.168
u
P   4.338, 1.949, -1.187
d
P   4.338, -0.879, 1.441
F
S
```

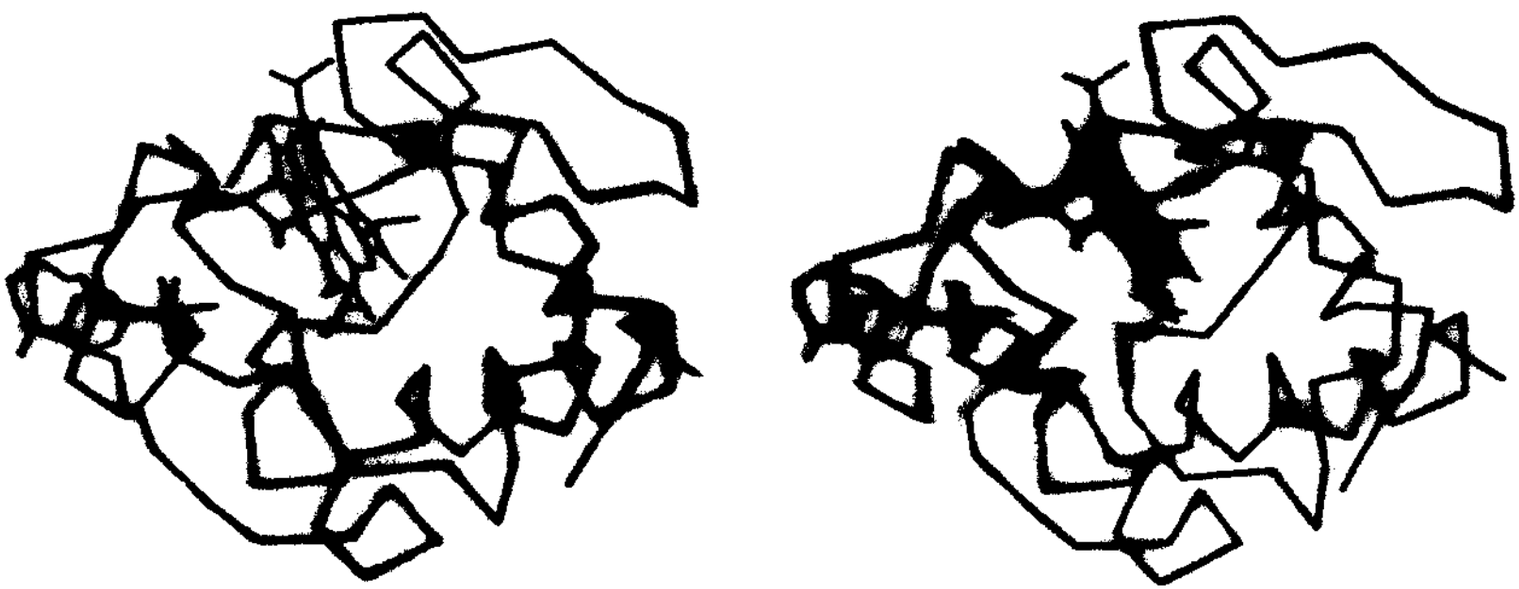
N1	2.3652,	0.3042,	-0.0041
C2	1.5073,	1.3262,	-0.0071
C3	0.1378,	1.1407,	0.0025
C4	-0.3652,	-0.1544,	0.0025
C5	0.5239,	-1.2116,	0.0053
C6	1.8837,	-0.9447,	0.0061
C7	-1.8496,	-0.4087,	0.0057
O1	-2.5626,	0.6707,	-0.0310
O2	-2.2939,	-1.5384,	0.0519
Ru	4.338,	0.535,	0.227
A	4.311,	0.766,	0.458
B	2.365,	0.304,	-0.004
C	4.012,	1.930,	1.422
D	4.645,	-0.860,	-1.168
E	4.338,	1.949,	-1.187
F	4.338,	-0.879,	1.641

I.3.2. Small Molecule Attachment

Using the files created in I.3.1, we can now display and orient the Ru - isonicotinate complex with cytochrome c. DISPLAY is used to create a protein entity (see section I.2.2.1) containing the alpha carbon backbone, lysine 72 side chain, and the heme prosthetic group. This is then displayed together with the Ru complex and the following operations performed: first an absolute translation of the Ru complex by the negatives of the coordinates of O1 which will be the attachment point. This keeps the center of rotation at the origin while translating O1 to the origin. Now a relative translation of the Ru complex by the coordinates of LYS 72 NZ as read from the Brookhaven file places O1 at the point of attachment. However, the cytochrome c has been translated from its original coordinates already by DISPLAY to place its center of rotation at the world coordinate system origin. The Ru complex must now be translated by this same relative amount. The translation is obtained by using the o command to activate 'cyt' and then using the t? command to query the protein's translation. The Ru complex is then made active again and a t command executed with the values obtained above. The results of these manipulations is shown in figure 27.

These manipulations have assumed the C7 to O1 bond length stays the same when O1 becomes NZ of lysine 72. This approximation is not unreasonable for visual purposes. The Ru complex may now be oriented into a sterically likely position. This

Figure 27: Stereo view of the Ru - isonicotinate complex attached to the terminal N of LYS 72 in cytochrome *c*.



is best done by viewing in three dimensions on the screen and experimenting with various angles about the attachment point. When a likely orientation is achieved, a global rotation may be applied to see the whole macromolecule from a different angle. The result of such manipulations is shown in figure 28.

I.3.3. Intermolecular and Intramolecular Distance Calculations

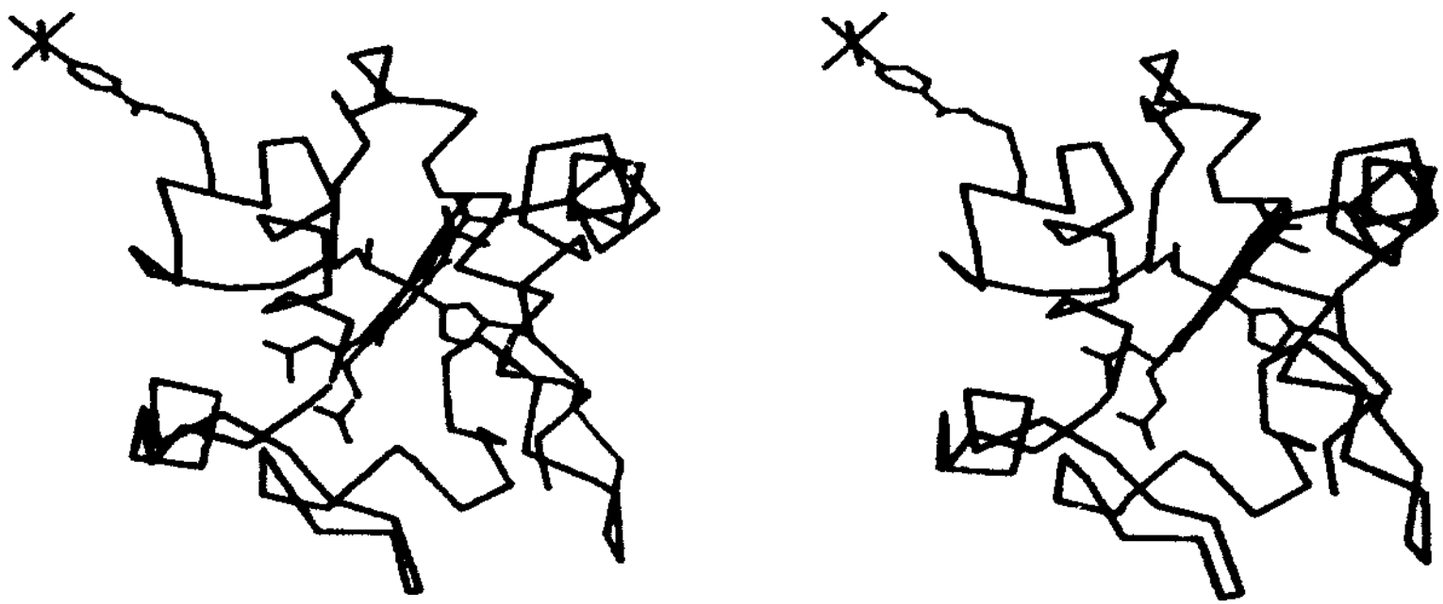
The example developed above may now be used to determine various distances by following the method of section I.2.3.3. Distances of interest may include the Ru - Fe distance, or Ru to lysine terminal N distances. A range of likely radii may be determined by trying different orientations which seem to be visually probable conformations.

I.3.4. Line of Sight Information

An interesting question to ask would be what aromatic side chains are along the linear path from Ru to Fe. Such information is easily obtainable by creating a protein entity with only a backbone and aromatic side chains and then orienting the whole complex so the Ru - Fe line is slighty skewed from the viewer's z-axis. This type of information would be extremely tedious to obtain from pure crystallographic information.

The necessary orientation is achieved by first attaching the Ru complex to the cytochrome lysine 72 as above, but with no global rotation applied. Then both the cytochrome

Figure 26: Stereo view of the Ru - isonicotinate complex attached to the terminal N of LYS 72 in cytochrome c and oriented into a sterically more likely position.



entity and Ru complex are relatively translated by the negatives of the original translation applied to the cytochrome entity by DISPLAY. Now both are again relatively translated by the negatives of the Fe coordinates as read from the Brookhaven file. This places the Fe at the world coordinates origin. A global rotation is then applied to put the Ru at a positive z-coordinate, slightly skewed from the z-axis.

A helpful thing to note about the example developed above is that the rotations and translations applied to the original Ru complex may be incorporated into the display command file for the Ru - isconicotinate. This will automatically orient the Ru complex to LYS 72 each time it is displayed with that particualr protein entity.

I.4. Source Code Layout

DISPLAY is organized into 20 source modules located on the SCS VAX at SYS$USER:[334003.3D]*.C. Each source file has a comment header with a brief description of the contents of that module. In addition, individual functions are internally commented where appropriate. Long function names and long variable names are used to augment the self commenting nature of the C language. The source code is generally organized into three levels of operation:

1) Basic graphics

2) Command interpreter and file interpreter

3) File creator and menus

The lowest level, basic graphics, is located in the following files:

MATRIX.C

VTGRAF.C

VTSET.C

VTTYPE.C

VTDECL.C

VTXTRN.C

VDEFN. C

Files VTDECL.C and VTXTRN.C contain global variables used by the graphics routines. VTDEFN.C contains definitions which may be changed if certain default values or device dependent

values are to be changed.

The next level contains functions which interpret display command files and call the basic graphics functions. The modules containing these functions are:

DISPDRAW.C

DISPEXEC.C

DISPINIT.C

DISPINP.C

DISPMOVE.C

DISPVIEW.C

DISPDECL.C

DISPXTRN.C

DISPDEFN.C

EXECXTRN.C

The highest level contains the functions which read a Brookhaven file and create display command files. They also specify active files, do distance calculations, and and present all menus. The modules containing these functions are:

DISPLAY.C

DISPRESI.C

DISPMANI.C

I.4.1. Device Independence

The definition file contains most of the characteristics of the particular physical device used (the Selanar Graphics - 100 board). If the definition file is changed,

the source modules using the changed definitions can be recompiled to operate with the new values. Most of the graphics code is device independent with the following functions being exceptions:

In VTGRAF.C -

1) vt_on ()

2) vt_on_text ()

3) vt_off ()

4) vt_clr ()

5) vt_move_physical ()

6) vt_draw_physical ()

7) vt_hair ()

With modifications of the definition file, VTDEFN.C, and changes to these seven functions, the DISPLAY program can be recompiled for use with another physical viewing device. Also, if ported to another computer, the system must have an equivalent to the System Plot Package (3) to produce metacode for separate hardcopy plotting. Subroutine calls to the System Plot Package occur in functions 1, 3, 5, 6 listed above and in the vt_plot_frame () function also in VTGRAF.C, and in function write_label () in DISPDRAW.C. These must be changed if DISPLAY is transported to another system.

I.4.2. Symbol Definitions for Compiling

Various symbols are defined in the file SYS$USER:[344003]LOGIN.COM. The symbols at the end of this file

starting with cc* allow easy recompilation and relinking of a specific module at a time. The symbols submit compiling command files and linking command files to the batch processor. The command files generally have the same name as the .C file but with a .COM extension.

I.4.3. Utility Symbol Definitions

SYS$USER:[344003]LOGIN.COM contains the following utility symbols:

cv to delete all DIS*.VT* files used by DISPLAY

cp to delete all IOP* plot files

cl to delete all .LOG files from batch compiling

mc to invoke the metacode translator

I.5. Miscellaneous Notes

This section contains pieces of useful information about the VAX and other items related to DISPLAY.

I.5.1. X-ray Department Programs

As mentioned before, programs are available to convert published crystallographic structural data into angstrom coordinates. Typing @SYS$SYSDEVICE:[XRAY]XSYMBOLS will define the correct symbols to run x-ray programs. Contact somebody in the x-ray department for information on how to use the programs. Sample .ATM and .INS files, which the user edits with structural information input to the x-ray programs, are available in SYS$USER:[334003.XRAY]. An example .FAX output file with angstrom coordinates for isonicotinic acid is also located in that directory.

I.5.2. Changing From FORTRAN Carriage Control

Files output from the x-ray programs (e.g. SYS$USER:[334003.XRAY]ISONIC.FAX) may be edited into small molecule coordinates files. However, after this is done, the carriage control attribute of the file must be changed from FORTRAN to carriage_return. To do this perform the following procedure:

```
$ ANALYZE/RMS/FDL FILENAME.EXT
```

```
$ EDIT/FDL/FILENAME    (no extension)
  choose the Modify option
  choose the Record attributes option
  choose carriage control attribute
  type "carriage" when prompted for the new attribute
  exit to the main menu and type "exit"
$ CONVERT/CREATE/FDL=FILENAME FILENAME.EXT FILENAME.EXT
```

I.5.3. The Metacode Translator

The metacode translator can be invoked properly for the Houston plotter by typing symbol mc (described in I.4.3). The metacode translator will interpret the most recently created (highest version number) metacode files (IOP*) and selectively allow plotting of metacode frames. Frames as used by DISPLAY are described in section I.2.2.1.

I.5.4. Brookhaven Files

Information on Brookhaven Protein Data Bank files may be tracked down through George Phillips at the University of Illinois.

I.5.5. Suggested Improvements and Known Bugs

As mentioned in I.2.3.2.2, it would be nice if global rotations could be implemented, but still allow additional local rotations to be applied relative to a coordinate system with the xy-plane being that of the screen with the z-axis coming forward to the viewer. Much of the code for this was

added to function rotate () in file DISPMOVE.C. Since it was not entirely completed by this writing, it has been commented out for now. Future work to implement this feature should be minimal. Notes on rotation transformations are available in Terry Coley's notes in the Scott group.

The comma separating parameters in commands such as the move to absolute point command (P) should not be necessary. Future work should modify the format strings in the scanf () functions invoked in the file interpreting functions.

At present, the line thickening (depth cueing) is determined by the line segment endpoint of highest viewing z coordinate. Thus, a line parallel to the viewing z-axis connected to a line parallel to the x or y-axis may give a corner with grossly mismatched thicknesses. A method of making the thickness at both ends of a line segment proportional to their corresponding z viewing coordinate should be developed. This involves modifying function vt_draw_physical () and will not be time comsuming.

One irregularity is known in the DISPLAY program at this time: DISPLAY prints an error message when reading the "Whole Chain" option of the side chain creator. This is due to the TER record and should not cause problems. See (6) for a description of the TER record.

Part II. Notes on the Ionic Strength Dependence of Cytochrome-c Redox Reactions with Small Molecules

Data has recently become available, through the work of James Schwartz in Dr. Robert Scott's group at the Univeristy of Illinois, for testing theories attempting to describe the ionic strength dependence of cytochrome c redox reactions with small molecules. Included in the data are rates of reactions with neutral iron complexes. Several theories have been put forth previously (1, 8 among others). Many make approximations which treat cytochrome c as a uniformly charged sphere of some effective radius. An attempt will not be made here to put forth a comparable theory. Rather, some possibly salient points will be made about what factors may be involved in a comprehensive theory. Notes on how the DISPLAY program may be used to help develop such a theory will be suggested.

The first point to make is that a general theory may require two parts. The first, which has been treated more extensively, is the long range effect of two charged particles in a dielectric medium. Work terms are computed and their contribution to the reaction rate estimated. The second part would include a short range effect where the dielectric medium may become inhomogeneous. For example a charged lysine may contribute to a long range charge effect, but have

less effect in the short range if it is located far from the reactive site with a lot of protein bulk in between.

In the short range a group of charged residues may create a potential well which directs the small molecule toward a reactive site. Such a possibility might be explored with the DISPLAY program by choosing a possible reaction site and using the distance calculator to obtain the positional and radial information to calculate a potential field. Such a local potential field would need to include a dielectric effect, which is ionic strength dependent, affecting the magnitude of the field.

Approximations of the dipole moment in cytochrome c have been made (10) and it has been suggested that the dipole helps orient the protein toward its approaching redox partner with an orientation favorable to reaction. Thus, a statistical orientation factor weighted by the effect of a dipole acting through a dielectric medium might be included. For a neutral partner, such a factor would reduce to a probability based solely on the percent of protein surface area which is reactive. The dipole orientation may be a good starting point for locating a small molecule using the DISPLAY program. The small molecule could then be visually oriented to bring it close to the protein while minimizing steric interactions. Then a potential field could be calculated for that area, possibly with an ionic strength dependent dielectric effect. The expected hydrophobicity of the site could suggest modifications of the dielectric effect, although this may

be hard to quantify.

One last consideration (which may or may not be applicable to Jim Schwartz's data) is the possibility of protein conformational changes induced by changes in ionic strength in solution. Such changes have been observed in certain systems (11) and other papers have more fully explored changes in redox potential with conformation in other proteins, such as cytochrome P-450(12) and ferric myoglobin hydroxide (13, 14). Changes in redox potential should have an effect on the rate of reaction and would be considered an indirect ionic strength dependence. Such conformational considerations would considerably complicate the theory.

References and Notes

(1) S. E. Wherland, Metalloprotein Electron Transfer Mechanisms, Thesis, California Institute of Technology, Pasadena, California (1977)

(2) B. W. Kernighan, D. M. Ritchie, The C Programming Language, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1978)

(3) G. R. McArthur, editor, The System Plot Package, NCAR Technical Note NCAR-TN/162+IA (1981)

(4) Houston Instrument HI-PLOT DMP-7 plotter, Houston Instrument, Austin, Texas 78753

(5) Selanar Graphics-100 Board, Selanar Corp., 437-A Aldo Ave., Santa Clara, California 95050

(6) Protein Data Bank Atomic Coordinate Entry Format Description, Dec 1981. Further information might be obtained from George Phillips at the University of Illinois

(7) J. D. Foley, A. VanDam, Fundamentals of Interactive Computer Graphics, Addison-Wesley Publishing Co., Reading, Massachusetts (1982)

(8) A. Shimada, F. Takisagawa, Acta Cryst., B32, 1925 (1976)

(9) R. P. Simondsen, P. C. Weber, F. R. Salemme, G. Tollin, Biochemistry, 21, 6366 (1982)

(10) M. I. Mitovska, K. I. Dancheva, B. P. Atanasov, Int. J. Biochem., 16, 1059 (1984)

(11) Goldkorn, Tzipora, Schejter, Abel, FEBS Lett., 75, 44 (1977)

(12) S. G. Sligar, I. C. Gunsalus, Biochemistry, 18, 2290 (1979)

(13) J. K. Beattie, R. J. West, JACS, 96, 1933 (1974)

(14) Dose, Tweedle, Wilson, JACS, 99, 3886 (1977)

(15) Genkin, Davydov, Krylov, Blumenfel'd, Biofizika, 22, 158 (1977)